

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Thomas, Terry L.

(ii) TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A  
DELTA 6-DESATURASE

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Scully, Scott, Murphy & Presser
- (B) STREET: 400 Garden City Plaza
- (C) CITY: Garden City
- (D) STATE: New York
- (E) COUNTRY: United States
- (F) ZIP: 11530

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Presser, Leopold
- (B) REGISTRATION NUMBER: 19,827
- (C) REFERENCE/DOCKET NUMBER: 8383ZYXWVU

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (516) 742-4343
- (B) TELEFAX: (516) 742-4366
- (C) TELEX: 230 901 SANS UR

10029755.122101

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3588 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2002..3081

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCTAGCCACC AGTGACGATG CCTTGAATTT GGCCATTCTG ACCCAGGCC	60
GTATTCTGAA	
TCCCGCATTT CGCATTTGTTA ATCGTTTGT	120
CAACCATGCC CTGGGTAAAC GTTTAGACAC	
CAGCTTGCCA GACCACGTTA GTTTGAGTGT TTCCGCCCTG	180
GCGGCCCCGA TTTTTCCTT	
TGCGGCTTTG GGCAATCAGG CGATCGGGCA ATTGCGTTTG	240
TTTGACCAGA CTTGGCCCAT	
TCAGGAAATT GTCAATTCACC AAGACCATCC CTGGCTCAAT	300
TTACCCCTGG CGGATTTATG	
GGATGATCCG AGCCGAATGT TGATCTATTA CCTACCGGCC	360
CACAGTGAAA CGGATTTAGT	
AGGCGCAGTG GTGAATAATT TAACGTTGCA ATCTGGGGAC	420
CATTTAATAG TGGGACAAAA	
ACCCCAACCC AAGACCAAAC GCGATCGCC TTGGCGCAA	480
TTTCCAAAC TGATTACCA	
CTGCGGGAG TATCAGCGGT ATGTCCAACA GGTGATATGG	540
GTGGTGTGT TTTTATTGTT	
GATGATTTTT CTGGCCACCT TCATCTACGT TTCCATTGAT	600
CAACATATTG CCCCAGTGGA	
CGCGTTGTAT TTTTCCGTGG GCATGATTAC CGGGCCGGT	660
GGCAAGGAAG AGGTGGCCGA	
AAAGTCCCC GATATCATCA AAGTATTCAC AGTGGTGATG	720
ATGATCGCCG GGC CGGGGGT	
GATTGGTATT TGTATGCCC TACTGAATGA TTTCATCCTT	780
GGCAGTCGCT TTAGTCAGTT	
TTTGATGCG GCCAAGTTAC CCGATCGCCA TCACATCATC	840
ATTGTGGG TGGGGGGAGT	
GAGCATGGCC ATTATTGAAG AGTTAATTCA CCAGGCCAT	900
GAAATTGTGG TAATCGAAAA	

GGATACAGAT AATCGTTTCT TGCATACGGC CCGCTCCCTG GGGGTGCCCG TAATTGTGGA	960
GGATGCCCCG CTAGAAAGAA CGTTGGCCTG CGCCAATATC AACCGAGCCG AAGCCATTGT	1020
GGTGCCACC AGCGACGACA CCGTTAACTT GGAAATTGGC CTAAGTCCA AGGCGATCGC	1080
CCCTAGCCCTG CCAGTGGTGT TGCCTTGCCA GGATGCCAG TTTAGCCTGT CCCTGCAGGA	1140
AGTATTTGAA TTTGAAACGG TGCTTTGTCC GGCAGGAATG GCCACCTATT CCTTTGCGGC	1200
GGCGGCCCTG GGGGGCAAAA TTTTGGGCAA CGGCATGACC GATGATTTGC TGTGGGTAGC	1260
CCTAGCCACC TTAATCACTC CTAACCATCC CTTTGCCGAC CAATTGGTTA AAATTGCAGC	1320
CAAAAGTCT GATTTCGTTC CCCTCTATCT AGAACGGGGT GGCAAAACCA TCCATAGCTG	1380
GGAATTATTG GGTACCCATC TCGACTCTGG AGACGTGTTG TATTTAACCA TGCCCGCCAC	1440
TGCCCTAGAG CAACCTTTGGC GATCGCCCCG TGCCACTGCT GATCCTCTGG ACTCTTTTTT	1500
GGTTTAGCAT GGGGGGATGG AACTCTTGAC TCGGCCCAAT GGTGATCAAG AAAGAACGCT	1560
TTCTCTATGT TTAGTATTTT TAAGTTAACC AACAGCAGAG GATAACTTCC AAAAGAAATT	1620
AAGCTCAAAA AGTAGCAAAA TAAGTTTAAT TCATAACTGA GTTTTACTGC TAAACAGCGG	1680
TGCAAAAAAG TCAGATAAAA TAAAAGCTTC ACTTCGGTTT TATATTGTGA CCATGGTTCC	1740
CAGCATCTG CTCTAGGGAG TTTTCCGCT GCCTTTAGAG AGTATTTTCT CCAAGTCGGC	1800
TAACTCCCC ATTTTTAGGC AAAATCATAT ACAGACTATC CCAATATTGC CAGAGCTTTG	1860
ATGACTCACT GTAGAAGGCA GACTAAAATT CTAGCAATGG ACTCCAGTT GGAATAAAAT	1920
TTTAGTCTCC CCCGCGCTG GAGTTTTTTT GTAGTTAATG GCGGTATAAT GTGAAAGTTT	1980
TTTATCTATT TAAATTTATA A ATG CTA ACA GCG GAA AGA ATT AAA TTT ACC	2031
Met Leu Thr Ala Glu Arg Ile Lys Phe Thr	
1 5 10	
CAG AAA CGG GGG TTT CGT CGG GTA CTA AAC CAA CGG GTG GAT GCC TAC	2079
Gln Lys Arg Gly Phe Arg Arg Val Leu Asn Gln Arg Val Asp Ala Tyr	
15 20 25	

TTT GCC GAG CAT GGC CTG ACC CAA AGG GAT AAT CCC TCC ATG TAT CTG	2127
Phe Ala Glu His Gly Leu Thr Gln Arg Asp Asn Pro Ser Met Tyr Leu	
30 35 40	
AAA ACC CTG ATT ATT GTG CTC TGG TTG TTT TCC GCT TGG GCC TTT GTG	2175
Lys Thr Leu Ile Ile Val Leu Trp Leu Phe Ser Ala Trp Ala Phe Val	
45 50 55	
CTT TTT GCT CCA GTT ATT TTT CCG GTG CGC CTA CTG GGT TGT ATG GTT	2223
Leu Phe Ala Pro Val Ile Phe Pro Val Arg Leu Leu Gly Cys Met Val	
60 65 70	
TTG GCG ATC GCC TTG GCG GCC TTT TCC TTC AAT GTC GGC CAC GAT GCC	2271
Leu Ala Ile Ala Leu Ala Phe Ser Phe Asn Val Gly His Asp Ala	
75 80 85 90	
AAC CAC AAT GCC TAT TCC TCC AAT CCC CAC ATC AAC CGG GTT CTG GGC	2319
Asn His Asn Ala Tyr Ser Ser Asn Pro His Ile Asn Arg Val Leu Gly	
95 100 105	
ATG ACC TAC GAT TTT GTC GGG TTA TCT AGT TTT CTT TGG CGC TAT CGC	2367
Met Thr Tyr Asp Phe Val Gly Leu Ser Ser Phe Leu Trp Arg Tyr Arg	
110 115 120	
CAC AAC TAT TTG CAC CAC ACC TAC ACC AAT ATT CTT GGC CAT GAC GTG	2415
His Asn Tyr Leu His His Thr Tyr Thr Asn Ile Leu Gly His Asp Val	
125 130 135	
GAA ATC CAT GGA GAT GGC GCA GTA CGT ATG AGT CCT GAA CAA GAA CAT	2463
Glu Ile His Gly Asp Gly Ala Val Arg Met Ser Pro Glu Gln Glu His	
140 145 150	
GTT GGT ATT TAT CGT TTC CAG CAA TTT TAT ATT TGG GGT TTA TAT CTT	2511
Val Gly Ile Tyr Arg Phe Gln Gln Phe Tyr Ile Trp Gly Leu Tyr Leu	
155 160 165 170	
TTC ATT CCC TTT TAT TGG TTC CTC TAC GAT GTC TAC CTA GTG CTT AAT	2559
Phe Ile Pro Phe Tyr Trp Phe Leu Tyr Asp Val Tyr Leu Val Leu Asn	
175 180 185	
AAA GGC AAA TAT CAC GAC CAT AAA ATT CCT CCT TTC CAG CCC CTA GAA	2607
Lys Gly Lys Tyr His Asp His Lys Ile Pro Pro Phe Gln Pro Leu Glu	
190 195 200	
TTA GCT AGT TTG CTA GGG ATT AAG CTA TTA TGG CTC GGC TAC GTT TTC	2655
Leu Ala Ser Leu Leu Gly Ile Lys Leu Leu Trp Leu Gly Tyr Val Phe	
205 210 215	
GGC TTA CCT CTG GCT CTG GGC TTT TCC ATT CCT GAA GTA TTA ATT GGT	2703
Gly Leu Pro Leu Ala Leu Gly Phe Ser Ile Pro Glu Val Leu Ile Gly	
220 225 230	
GCT TCG GTA ACC TAT ATG ACC TAT GGC ATC GTG GTT TGC ACC ATC TTT	2751
Ala Ser Val Thr Tyr Met Thr Tyr Gly Ile Val Val Cys Thr Ile Phe	
235 240 245 250	

ATG CTG GCC CAT GTG TTG GAA TCA ACT GAA TTT CTC ACC CCC GAT GGT	2799
Met Leu Ala His Val Leu Glu Ser Thr Glu Phe Leu Thr Pro Asp Gly	
255 260 265	
GAA TCC GGT GCC ATT GAT GAC GAG TGG GCT ATT TGC CAA ATT CGT ACC	2847
Glu Ser Gly Ala Ile Asp Asp Glu Trp Ala Ile Cys Gln Ile Arg Thr	
270 275 280	
ACG GCC AAT TTT GCC ACC AAT AAT CCC TTT TGG AAC TGG TTT TGT GGC	2895
Thr Ala Asn Phe Ala Thr Asn Asn Pro Phe Trp Asn Trp Phe Cys Gly	
285 290 295	
GGT TTA AAT CAC CAA GTT ACC CAC CAT CTT TTC CCC AAT ATT TGT CAT	2943
Gly Leu Asn His Gln Val Thr His His Leu Phe Pro Asn Ile Cys His	
300 305 310	
ATT CAC TAT CCC CAA TTG GAA AAT ATT ATT AAG GAT GTT TGC CAA GAG	2991
Ile His Tyr Pro Gln Leu Glu Asn Ile Ile Lys Asp Val Cys Gln Glu	
315 320 325 330	
TTT GGT GTG GAA TAT AAA GTT TAT CCC ACC TTC AAA GCG GCG ATC GCC	3039
Phe Gly Val Glu Tyr Lys Val Tyr Pro Thr Phe Lys Ala Ala Ile Ala	
335 340 345	
TCT AAC TAT CGC TGG CTA GAG GCC ATG GGC AAA GCA TCG TGACATTGCC	3088
Ser Asn Tyr Arg Trp Leu Glu Ala Met Gly Lys Ala Ser	
350 355 360	
TTGGGATTGA AGCAAAATGG CAAAATCCCT CGTAAATCTA TGATCGAAGC CTTTCTGTTG	3148
CCCGCCGACC AAATCCCCGA TGCTGACCAA AGGTTGATGT TGGCATTGCT CCAAACCCAC	3208
TTTGAGGGGG TTCATTGGCC GCAGTTTCAA GCTGACCTAG GAGGCAAAGA TTGGGTGATT	3268
TTGCTCAAAT CCGCTGGGAT ATTGAAAGGC TTCACCACCT TTGGTTTCTA CCCTGCTCAA	3328
TGGGAAGGAC AAACCGTCAG AATTGTTTAT TCTGGTGACA CCATACCCGA CCCATCCATG	3388
TGCTCTAACC CAGCCCTGGC CAAGGCTTGG ACCAAGGCCA TGCAAATTCT CCACGAGGCT	3448
AGGCCAGAAA AATTATATTG GCTCCTGATT TCTTCCGGCT ATCGCACCTA CCGATTTTTG	3508
AGCATTTTTG CCAAGGAATT CTATCCCCAC TATCTCCATC CCACTCCCCC GCCTGTACAA	3568
AATTTTATCC ATCAGCTAGC	3588

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 359 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Leu	Thr	Ala	Glu	Arg	Ile	Lys	Phe	Thr	Gln	Lys	Arg	Gly	Phe	Arg
1				5					10					15	
Arg	Val	Leu	Asn	Gln	Arg	Val	Asp	Ala	Tyr	Phe	Ala	Glu	His	Gly	Leu
		20						25					30		
Thr	Gln	Arg	Asp	Asn	Pro	Ser	Met	Tyr	Leu	Lys	Thr	Leu	Ile	Ile	Val
		35					40					45			
Leu	Trp	Leu	Phe	Ser	Ala	Trp	Ala	Phe	Val	Leu	Phe	Ala	Pro	Val	Ile
	50					55					60				
Phe	Pro	Val	Arg	Leu	Leu	Gly	Cys	Met	Val	Leu	Ala	Ile	Ala	Leu	Ala
	65				70					75					80
Ala	Phe	Ser	Phe	Asn	Val	Gly	His	Asp	Ala	Asn	His	Asn	Ala	Tyr	Ser
				85					90					95	
Ser	Asn	Pro	His	Ile	Asn	Arg	Val	Leu	Gly	Met	Thr	Tyr	Asp	Phe	Val
			100					105					110		
Gly	Leu	Ser	Ser	Phe	Leu	Trp	Arg	Tyr	Arg	His	Asn	Tyr	Leu	His	His
		115					120					125			
Thr	Tyr	Thr	Asn	Ile	Leu	Gly	His	Asp	Val	Glu	Ile	His	Gly	Asp	Gly
	130					135					140				
Ala	Val	Arg	Met	Ser	Pro	Glu	Gln	Glu	His	Val	Gly	Ile	Tyr	Arg	Phe
	145				150					155					160
Gln	Gln	Phe	Tyr	Ile	Trp	Gly	Leu	Tyr	Leu	Phe	Ile	Pro	Phe	Tyr	Trp
				165					170					175	
Phe	Leu	Tyr	Asp	Val	Tyr	Leu	Val	Leu	Asn	Lys	Gly	Lys	Tyr	His	Asp
			180					185					190		
His	Lys	Ile	Pro	Pro	Phe	Gln	Pro	Leu	Glu	Leu	Ala	Ser	Leu	Leu	Gly
		195					200					205			
Ile	Lys	Leu	Leu	Trp	Leu	Gly	Tyr	Val	Phe	Gly	Leu	Pro	Leu	Ala	Leu
	210					215					220				
Gly	Phe	Ser	Ile	Pro	Glu	Val	Leu	Ile	Gly	Ala	Ser	Val	Thr	Tyr	Met
	225				230					235					240
Thr	Tyr	Gly	Ile	Val	Val	Cys	Thr	Ile	Phe	Met	Leu	Ala	His	Val	Leu
			245						250					255	
Glu	Ser	Thr	Glu	Phe	Leu	Thr	Pro	Asp	Gly	Glu	Ser	Gly	Ala	Ile	Asp
			260					265						270	

Asp Glu Trp Ala Ile Cys Gln Ile Arg Thr Thr Ala Asn Phe Ala Thr  
 275 280 285  
 Asn Asn Pro Phe Trp Asn Trp Phe Cys Gly Gly Leu Asn His Gln Val  
 290 295 300  
 Thr His His Leu Phe Pro Asn Ile Cys His Ile His Tyr Pro Gln Leu  
 305 310 315 320  
 Glu Asn Ile Ile Lys Asp Val Cys Gln Glu Phe Gly Val Glu Tyr Lys  
 325 330 335  
 Val Tyr Pro Thr Phe Lys Ala Ala Ile Ala Ser Asn Tyr Arg Trp Leu  
 340 345 350  
 Glu Ala Met Gly Lys Ala Ser  
 355

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1884 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCTTCACTT CGGTTTTATA TTGTGACCAT GGTTCACAGG CATCTGCTCT AGGGAGTTT	60
TCCGCTGCCT TTAGAGAGTA TTTTCTCCAA GTCGGCTAAC TCCCCATTT TTAGGCAAAA	120
TCATATACAG ACTATCCCAA TATTGCCAGA GCTTTGATGA CTCACTGTAG AAGGCAGACT	180
AAAATTCTAG CAATGGACTC CCAGTTGGAA TAAATTTTTA GTCTCCCCCG GCGCTGGAGT	240
TTTTTTGTAG TTAATGGCGG TATAATGTGA AAGTTTTTTA TCTATTAA TTTATAAATG	300
CTAACAGCGG AAAGAATTAA ATTTACCCAG AAACGGGGGT TTCGTCGGGT ACTAAACCAA	360
CGGGTGGATG CCTACTTTGC CGAGCATGGC CTGACCCAAA GGGATAATCC CTCCATGTAT	420
CTGAAAACCC TGATTATTGT GCTCTGGTTG TTTTCCGCTT GGGCCTTTGT GCTTTTGTCT	480
CCAGTTATTT TTCCGGTGCG CCTACTGGGT TGTATGGTTT TGGCGATCGC CTTGGCGGCC	540
TTTTCCTTCA ATGTCGGCCA CGATGCCAAC CACAATGCCT ATTCCTCCAA TCCCCACATC	600
AACCGGGTTC TGGGCATGAC CTACGATTTT GTCGGGTAT CTAGTTTCT TTTGGCGCTAT	660
CGCCACAAC TTTTGCACCA CACCTACACC AATATTCTTG GCCATGACGT GGAAATCCAT	720

GGAGATGGCG	CAGTACGTAT	GAGTCTTGAA	CAAGAACATG	TTGGTATTTA	TCGTTTCCAG	780
CAATTTTATA	TTTGGGGTTT	ATATCTTTTC	ATTCCTTTT	ATTGGTTTCT	CTACGATGTC	840
TACCTAGTGC	TTAATAAAGG	CAATATACAC	GACCATAAAA	TTCTCCTTT	CCAGCCCCTA	900
GAATTAGCTA	GTTTGCTAGG	GATTAAGCTA	TTATGGCTCG	GCTACGTTTT	CGGCTTACCT	960
CTGGCTCTGG	GCTTTTCCAT	TCCTGAAGTA	TTAATTGGTG	CTTCGGTAAC	CTATATGACC	1020
TATGGCATCG	TGGTTTGCAC	CATCTTTATG	CTGGCCCATG	TGTTGGAATC	AACTGAATTT	1080
CTCACCCCCG	ATGGTGAATC	CGGTGCCATT	GATGACGAGT	GGGCTATTTG	CCAAATTCGT	1140
ACCACGGCCA	ATTTTGGCAC	CAATAATCCC	TTTTGGAAT	GGTTTGTGTG	CGGTTTAAAT	1200
CACCAAGTTA	CCCACCATCT	TTTCCCCAAT	ATTTGTCTA	TTCATATCC	CCAATTGGAA	1260
AATATTATTA	AGGATGTTTG	CCAAGAGTTT	GGTGTGGAAT	ATAAAGTTTA	TCCACCTTC	1320
AAACCGGCGA	TCGCTCTAA	CTATCGCTGG	CTAGAGGCCA	TGGGCAAAGC	ATCGTGACAT	1380
TGCTTTGGGA	TTGAAGCAAA	ATGGCAAAAT	CCCTCGTAAA	TCTATGATCG	AAGCCTTTCT	1440
GTTCGCCGCC	GACCAAAATC	CCGATGCTGA	CCAAAGGTTG	ATGTTGGCAT	TGCTCCAAAC	1500
CCACTTTGAG	GGGGTTCATT	GGCCGCAAGT	TCAAGCTGAC	CTAGGAGGCA	AAGATTGGGT	1560
GATTTTGCTC	AAATCCGCTG	GGATATTGAA	AGGCTTCACC	ACCTTTGGTT	TCTACCCTGC	1620
TCAATGGGAA	GGACAAACCG	TCAGAATTGT	TTATTTCTGGT	GACACCATCA	CCGACCCATC	1680
CATCTGGTCT	AACCCAGCCC	TGGCCAAGGC	TTGGACCAAG	GCCATGCAAA	TTCTCCACGA	1740
GGCTAGGCCA	GAATAATTAT	ATTGGCTCCT	GATTCTTTCC	GGCTATCGCA	CCTACCGATT	1800
TTTGAGCATT	TTTGCCAAGG	AATTCTATCC	CCACTATCTC	CATCCCACTC	CCCCGCCTGT	1860
ACAAAATTTT	ATCCATCAGC	TAGC				1884

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1685 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATATCTGCC TACCTCCCA AAGAGAGTAG TCATTTTCA TCAATGGCTG CTCAAATCAA



GAAATACATT	ACCTCAGATG	AACTCAAGAA	CCACGATAAA	CCCGGAGATC	TATGGATCTC	120
GATTC AAGGG	AAAGCCTATG	ATGTTTCGGA	TTGGGTGAAA	GACCATCCAG	GTGGCAGCTT	180
TCCCTTGAAG	AGTCTTGCTG	GTC AAGAGGT	AACTGATGCA	TTTGTTCGAT	TCCATCCTGC	240
CTCTACATGG	AAGAATCTTG	ATAAGTTTIT	CACTGGGTAT	TATCTTAAAG	ATTACTCTGT	300
TTCTGAGGTT	TCTAAAGATT	ATAGGAAGCT	TGTGTTTGAG	TTTTCTAAAA	TGGGTTTGTA	360
TGACAAAAAA	GGTCATATTA	TGTTTGCAAC	TTTGTGCTTT	ATAGCAATGC	TGTTTGCTAT	420
GAGTGTITAT	GGGGTTTTGT	TTTGTGAGGG	TGTTTTGGTA	CATTTGTITTT	CTGGGTGTTT	480
GATGGGGTTT	CTTTGGATTG	AGAGTGGTTG	GATTGGACAT	GATGCTGGGC	ATTATATGGT	540
AGTGCTCTGAT	TCAAGGCTTA	ATAAGTTTAT	GGGTATTTIT	GCTGCAAAAT	GTCTTTTCAGG	600
AAATAAGTATT	GGTTGGTGGA	AATGGAACCA	TAATGCACAT	CACATTGCCT	GTAATAGCCT	660
TGAATATGAC	CCTGATTTAC	AATATATACC	ATTCTCTGTT	TGTGCTTCCA	AGTTTTTTTG	720
TTCACTCACC	TCTCATTTCT	ATGAGAAAAAG	GTTGACTTTT	GACTCTTTAT	CAAGATTCCT	780
TGTAAGTTAT	CAACATTGGA	CATTTTACCC	TATTATGTGT	GCTGCTAGGC	TCAATATGTA	840
TGTACAATCT	CTCATAATGT	TGTTGACCAA	GAGAAATGTG	TCCTATCGAG	CTCAGGAAC	900
CTTGGGATGC	CTAGTGTCT	CGATTGGTA	CCCGTTGCTT	GTTTCTTGTT	TGCCTAATTG	960
GGGTGAAAGA	ATTATGTTTG	TTATTGCAAG	TTTATCAGTG	ACTGGAATGC	AACAAGTTCA	1020
GTCTCTCTTG	AACCACTTCT	CTTCAAGTGT	TTATGTTGGA	AAGCCTAAAG	GGAATAATTG	1080
GTTTGAGAAA	CAAACGGATG	GGACACTTGA	CATTTCTTGT	CCTCCTTGA	TGGATTGGTT	1140
TCATGGTGGA	TTGCAATTCC	AAATTGAGCA	TCATTTGTTT	CCCAAGATGC	CTAGATGCAA	1200
CCTTAGGAAA	ATCTCGCCCT	ACGTGATCGA	GTTATGCAAG	AAACATAAAT	TGCCTTTACAA	1260
TTATGCATCT	TTCTCCAAGG	CCAATGAAAT	GACACTCAGA	ACATTGAGGA	ACACAGCATT	1320
GCAGGCTAGG	GATATAACCA	AGCCGCTCCC	GAAGAATTTG	GTATGGGAAG	CTCTTCACAC	1380
TCATGGTTAA	AATTACCCTT	AGTTTCATGTA	ATAATTTGAG	ATTATGTATC	TCCTATGTTT	1440
GTGTCTTGTC	TTGGTTCTAC	TGTTGGAGT	CATTGCAACT	TGTCTTTTAT	GGTTTATTAG	1500
ATGTTTTTTA	ATATATTTTA	GAGGTTTTGC	TTTCATCTCC	ATTATTGATG	AATAAGGAGT	1560
TGCATATTGT	CAATTGTTGT	GCTCAATATC	TGATATTTTG	GAATGTACTT	TGTACCACTG	1620
TGTTTTTCAGT	TGAAGCTCAT	GTGTACTTCT	ATAGACTTTG	TTTAAATGGT	TATGTCATGT	1680
TATTTT						1685

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 448 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ala	Ala	Gln	Ile	Lys	Lys	Tyr	Ile	Thr	Ser	Asp	Glu	Leu	Lys	Asn	1	5	10	15
His	Asp	Lys	Pro	Gly	Asp	Leu	Trp	Ile	Ser	Ile	Gln	Gly	Lys	Ala	Tyr	20	25	30	
Asp	Val	Ser	Asp	Trp	Val	Lys	Asp	His	Pro	Gly	Gly	Ser	Phe	Pro	Leu	35	40	45	
Lys	Ser	Leu	Ala	Gly	Gln	Glu	Val	Thr	Asp	Ala	Phe	Val	Ala	Phe	His	50	55	60	
Pro	Ala	Ser	Thr	Trp	Lys	Asn	Leu	Asp	Lys	Phe	Phe	Thr	Gly	Tyr	Tyr	65	70	75	80
Leu	Lys	Asp	Tyr	Ser	Val	Ser	Glu	Val	Ser	Lys	Asp	Tyr	Arg	Lys	Leu	85	90	95	
Val	Phe	Glu	Phe	Ser	Lys	Met	Gly	Leu	Tyr	Asp	Lys	Lys	Gly	His	Ile	100	105	110	
Met	Phe	Ala	Thr	Leu	Cys	Phe	Ile	Ala	Met	Leu	Phe	Ala	Met	Ser	Val	115	120	125	
Tyr	Gly	Val	Leu	Phe	Cys	Glu	Gly	Val	Leu	Val	His	Leu	Phe	Ser	Gly	130	135	140	
Cys	Leu	Met	Gly	Phe	Leu	Trp	Ile	Gln	Ser	Gly	Trp	Ile	Gly	His	Asp	145	150	155	160
Ala	Gly	His	Tyr	Met	Val	Val	Ser	Asp	Ser	Arg	Leu	Asn	Lys	Phe	Met	165	170	175	
Gly	Ile	Phe	Ala	Ala	Asn	Cys	Leu	Ser	Gly	Ile	Ser	Ile	Gly	Trp	Trp	180	185	190	
Lys	Trp	Asn	His	Asn	Ala	His	His	Ile	Ala	Cys	Asn	Ser	Leu	Glu	Tyr	195	200	205	
Asp	Pro	Asp	Leu	Gln	Tyr	Ile	Pro	Phe	Leu	Val	Val	Ser	Ser	Lys	Phe	210	215	220	

10027556.1.22101

Phe Gly Ser Leu Thr Ser His Phe Tyr Glu Lys Arg Leu Thr Phe Asp  
 225 230 235 240  
 Ser Leu Ser Arg Phe Phe Val Ser Tyr Gln His Trp Thr Phe Tyr Pro  
 245 250 255  
 Ile Met Cys Ala Ala Arg Leu Asn Met Tyr Val Gln Ser Leu Ile Met  
 260 265 270  
 Leu Leu Thr Lys Arg Asn Val Ser Tyr Arg Ala Gln Glu Leu Leu Gly  
 275 280 285  
 Cys Leu Val Phe Ser Ile Trp Tyr Pro Leu Leu Val Ser Cys Leu Pro  
 290 295 300  
 Asn Trp Gly Glu Arg Ile Met Phe Val Ile Ala Ser Leu Ser Val Thr  
 305 310 315 320  
 Gly Met Gln Gln Val Gln Phe Ser Leu Asn His Phe Ser Ser Ser Val  
 325 330 335  
 Tyr Val Gly Lys Pro Lys Gly Asn Asn Trp Phe Glu Lys Gln Thr Asp  
 340 345 350  
 Gly Thr Leu Asp Ile Ser Cys Pro Pro Trp Met Asp Trp Phe His Gly  
 355 360 365  
 Gly Ser Gln Phe Gln Ile Glu His His Leu Phe Pro Lys Met Pro Arg  
 370 375 380  
 Cys Asn Leu Arg Lys Ile Ser Pro Tyr Val Ile Glu Leu Cys Lys Lys  
 385 390 395 400  
 His Asn Leu Pro Tyr Asn Tyr Ala Ser Phe Ser Lys Ala Asn Glu Met  
 405 410 415  
 Thr Leu Arg Thr Leu Arg Asn Thr Ala Leu Gln Ala Arg Asp Ile Thr  
 420 425 430  
 Lys Pro Leu Pro Lys Asn Leu Val Trp Glu Ala Leu His Thr His Gly  
 435 440 445

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

10029755.122101

Trp Ile Gly His Asp Ala Gly His  
1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Asn Val Gly His Asp Ala Asn His  
1 5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Val Leu Gly His Asp Cys Gly His  
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (peptide)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Val Ile Ala His Glu Cys Gly His  
1 5

10029756.12101

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Val Ile Gly His Asp Cys Ala His  
1 5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Val Val Gly His Asp Cys Gly His  
1 5

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

His Asn Ala His His  
1 5

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (peptide)

10029756.122101

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

His Asn Tyr Leu His His  
1 5

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

His Arg Thr His His  
1 5

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

His Arg Arg His His  
1 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

His Asp Arg His His  
1 5

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

His Asp Gln His His  
1 5

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

His Asp His His His  
1 5

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

His Asn His His His  
1 5

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

10029755-125701

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Phe Gln Ile Glu His His  
1 5

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

His Gln Val Thr His His  
1 5

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

His Val Ile His His  
1 5

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

His Val Ala His His  
1 5

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids



(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

His Ile Pro His His  
1 5

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

His Val Pro His His  
1 5

2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1702 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 48..1406

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 48..1406

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CCCCAAAAAT TTTCATTGTT CTCCATCTGG ACCACAGCAT CCACACA ATG GAG GGC	56
Met Glu Gly	
1	
GAA GCT AAG AAG TAT ATC ACG GCG GAG GAC CTC CGC CGC CAC AAC AAG	104
Glu Ala Lys Lys Tyr Ile Thr Ala Glu Asp Leu Arg Arg His Asn Lys	
5 10 15	
TCC GGC GAT CTC TGG ATC TCC ATC CAG GGC AAG GTC TAC GAC TGC TCT	152

Ser 20	Gly	Asp	Leu	Trp	Ile 25	Ser	Ile	Gln	Gly 30	Lys	Val	Tyr	Asp	Cys	Ser 35	
CGG	TGG	GCG	GCG	GAG	CAC	CCC	GGC	GGC	GAG	GTC	CCG	CTC	CTC	AGT	CTG	200
Arg	Trp	Ala	Ala	Glu 40	His	Pro	Gly	Gly	Glu 45	Val	Pro	Leu	Leu	Ser 50	Leu	
GCC	GCG	CAG	GAC	GTC	ACC	GAC	GCC	TTC	ATT	GCG	TAC	CAC	CCG	GGC	ACG	248
Ala	Gly	Gln	Asp 55	Val	Thr	Asp	Ala	Phe 60	Ile	Ala	Tyr	His	Pro 65	Gly	Thr	
GCG	TGG	CGG	CAT	CTG	GAT	CCG	CTC	TTC	ACC	GGC	TAC	TAC	TAC	CTC	AAG	296
Ala	Trp	Arg 70	His	Leu	Asp	Pro	Leu 75	Phe	Thr	Gly	Tyr	Tyr 80	Tyr	Leu	Lys	
GAC	TTC	GAA	GTG	TCG	GAG	ATC	TCC	AAG	GAC	TAC	CGG	AGG	CTT	TTG	AAC	344
Asp	Phe	Glu 85	Val	Ser	Glu 90	Ile	Ser	Lys	Asp	Tyr	Arg 95	Arg	Leu	Leu	Asn	
GAG	ATG	TCG	CGG	TCC	GGG	ATC	TTC	GAG	AAG	AAG	GGC	CAC	CAC	ATC	ATG	392
Glu 100	Met	Ser	Arg	Ser 105	Gly	Ile	Phe	Glu	Lys 110	Lys	Gly	His	His	Ile 115	Met	
TGG	ACG	TTC	GTC	GGC	GTT	GCG	GTC	ATG	ATG	GCG	GCA	ATC	GTC	TAC	GGC	440
Trp	Thr	Phe	Val 120	Gly	Val	Ala	Val	Met	Met 125	Ala	Ala	Ile	Val	Tyr 130	Gly	
GTG	CTG	GCG	TCG	GAG	TCC	GTC	GGA	GTT	CAC	ATG	CTC	TGC	GGC	GCA	CTG	488
Val	Leu	Ala	Ser 135	Glu	Ser	Val	Gly 140	Val	His	Met	Leu	Cys	Gly 145	Ala	Leu	
CTG	GGC	TTG	CTG	TGG	ATC	CAA	GCC	GCG	TAT	GTG	GGC	CAT	GAC	TCC	GGC	536
Leu	Gly	Leu 150	Leu	Trp	Ile	Gln	Ala 155	Ala	Tyr	Val	Gly	His 160	Asp	Ser	Gly	
CAT	TAC	CAG	GTG	ATG	CCA	ACC	CGT	GGA	TAC	AAC	AGA	ATC	ACG	CAA	CTC	584
His	Tyr 165	Gln	Val	Met	Pro 170	Thr	Arg	Gly	Tyr	Asn 175	Arg	Ile	Thr	Gln	Leu	
ATA	GCA	GGC	AAC	ATC	CTA	ACC	GGA	ATC	AGC	ATC	GCG	TGG	TGG	AAG	TGG	632
Ile	Ala	Gly	Asn	Ile	Leu 185	Thr	Gly	Ile	Ser	Ile 190	Ala	Trp	Trp	Lys 195	Trp	
ACC	CAC	AAC	GCC	CAC	CAC	CTC	GCC	TGC	AAC	AGC	CTC	GAC	TAC	GAC	CCC	680
Thr	His	Asn	Ala 200	His	His	Leu	Ala	Cys	Asn 205	Ser	Leu	Asp	Tyr	Asp 210	Pro	
GAC	CTC	CAG	CAC	ATC	CCC	GTA	TTC	GCC	GTC	TCC	ACC	CGA	CTC	TTC	AAC	728
Asp	Leu	Gln 215	His	Ile	Pro	Val	Phe 220	Ala	Val	Ser	Thr	Arg	Leu 225	Phe	Asn	
TCC	ATC	ACC	TCG	GTC	TTC	TAT	GGC	CGA	GTC	CTG	AAA	TTC	GAC	GAA	GTG	776
Ser	Ile	Thr 230	Ser	Val	Phe	Tyr	Gly 235	Arg	Val	Leu	Lys	Phe 240	Asp	Glu	Val	

GCA Ala	CGG Arg	TTC Phe	CTA Leu	GTC Val	AGC Ser	TAC Tyr	CAG Gln	CAC His	TGG Trp	ACC Thr	TAC Tyr	CCG Pro	GTC Val	ATG Met	824
ATC Ile	TTC Phe	GGC Gly	CGA Arg	GTC Val	AAC Asn	CTC Leu	TTC Phe	ATC Ile	CAG Gln	ACC Thr	TTT Phe	TTA Leu	TTG Leu	CTC Leu	872
ACC Thr	AGG Arg	CGC Arg	GAC Asp	GTC Val	CCT Pro	GAC Asp	CGC Arg	GCT Ala	CTA Leu	AAC Asn	TTA Leu	ATG Met	GGT Gly	ATC Ile	920
GTT Val	TTC Phe	TGG Trp	ACG Thr	TGG Trp	TTC Phe	CCG Pro	CTC Leu	TTC Phe	GTA Val	TCT Ser	TGT Cys	CTC Leu	CCG Pro	AAC Asn	968
CCT Pro	GAA Glu	CGG Arg	TTC Phe	GGG Gly	TTC Phe	GTC Val	CTC Leu	ATC Ile	AGC Ser	TTT Phe	GCG Ala	GTC Val	ACG Thr	GCG Ala	1016
CAG Gln	CAC His	GTC Val	CAG Gln	TTC Phe	ACG Thr	CTC Leu	AAC Asn	CAC His	TTC Phe	TCC Ser	GGC Gly	GAC Asp	ACA Thr	TAC Tyr	1064
GGC Gly	CCC Pro	CCC Pro	AAG Lys	GGC Gly	GAC Asp	AAC Asn	TGG Trp	TTC Phe	GAG Glu	AAG Lys	CAG Gln	ACG Thr	AAA Lys	GGG Gly	1112
ATC Ile	GAT Asp	ATC Ile	ACG Thr	TGC Cys	CCA Pro	CCG Pro	TGG Trp	ATG Met	GAC Asp	TGG Trp	TTC Phe	TTT Phe	GGT Gly	GGG Gly	1160
CAG Gln	TTC Phe	CAG Gln	TTG Leu	GAG Glu	CAC His	CAC His	TTG Leu	TTC Phe	CCT Pro	AGG Arg	CTG Leu	CCG Pro	CGT Arg	GGG Gly	1208
CTT Leu	AGG Arg	AAG Lys	ATT Ile	GCG Ala	CCC Pro	TTG Leu	GCT Ala	CGG Arg	GAC Asp	TTG Leu	TGT Cys	AAG Lys	AAG Lys	CAC His	1256
ATG Met	CCG Pro	TAT Tyr	AGG Arg	AGC Ser	TTC Phe	GGG Gly	TTT Phe	TGG Trp	GAC Asp	GAC Asp	GCT Ala	AAT Asn	GTC Val	AGG Arg	1304
ATT Ile	CGG Arg	ACG Thr	CTG Leu	AGG Arg	GAT Asp	GCG Ala	GCG Ala	GTT Val	CAG Gln	GCG Ala	CGT Arg	GAC Asp	CTT Leu	AAT Asn	1352
GCC Ala	CCG Pro	TGC Cys	CCT Pro	AAG Lys	AAA Lys	CTT Leu	GGG Gly	TAT Tyr	GGG Gly	GAA Glu	GCT Ala	TAT Tyr	AAC Asn	ACC Thr	1400
GGT Gly	TGA *	TTGTGGT	TTTT	GTGTTGTG	TGGG	TTGGAGG	GATC	TTCTTATT	TAT	TGATT	TATGT				1456

CCACAATATT GAACTGAATA ACCATGGAAG GCACTACGTT CAGCTTAACT TTGCTTAACT	1516
TGTCTAGCTG GTTGCCTTCC CTTGTTGGGG GCAAAGTGCA GTATTTATTT TCTTATCCCA	1576
TGTACTTTTT GATTATTGTT CTTATTCGTA TCATAAATAA TTTATTATTG ATTAATTTTT	1636
GTTGTAGTTG GGTGTCTATA GCAAGTTTAT AATACTGAGA TATATTTTTT TGGTAAAAAA	1696
AAAAAA	1702

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 453 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

1000715-12201

Met	Glu	Gly	Glu	Ala	Lys	Lys	Tyr	Ile	Thr	Ala	Glu	Asp	Leu	Arg	Arg
1				5					10					15	
His	Asn	Lys	Ser	Gly	Asp	Leu	Trp	Ile	Ser	Ile	Gln	Gly	Lys	Val	Tyr
			20					25					30		
Asp	Cys	Ser	Arg	Trp	Ala	Ala	Glu	His	Pro	Gly	Gly	Glu	Val	Pro	Leu
		35					40					45			
Leu	Ser	Leu	Ala	Gly	Gln	Asp	Val	Thr	Asp	Ala	Phe	Ile	Ala	Tyr	His
	50					55					60				
Pro	Gly	Thr	Ala	Trp	Arg	His	Leu	Asp	Pro	Leu	Phe	Thr	Gly	Tyr	Tyr
	65				70					75					80
Tyr	Leu	Lys	Asp	Phe	Glu	Val	Ser	Glu	Ile	Ser	Lys	Asp	Tyr	Arg	Arg
				85					90					95	
Leu	Leu	Asn	Glu	Met	Ser	Arg	Ser	Gly	Ile	Phe	Glu	Lys	Lys	Gly	His
			100					105					110		
His	Ile	Met	Trp	Thr	Phe	Val	Gly	Val	Ala	Val	Met	Met	Ala	Ala	Ile
		115					120					125			
Val	Tyr	Gly	Val	Leu	Ala	Ser	Glu	Ser	Val	Gly	Val	His	Met	Leu	Cys
	130					135					140				
Gly	Ala	Leu	Leu	Gly	Leu	Leu	Trp	Ile	Gln	Ala	Ala	Tyr	Val	Gly	His
	145				150					155					160
Asp	Ser	Gly	His	Tyr	Gln	Val	Met	Pro	Thr	Arg	Gly	Tyr	Asn	Arg	Ile
				165					170					175	

Thr Gln Leu Ile Ala Gly Asn Ile Leu Thr Gly Ile Ser Ile Ala Trp  
180 185 190

Trp Lys Trp Thr His Asn Ala His His Leu Ala Cys Asn Ser Leu Asp  
195 200 205

Tyr Asp Pro Asp Leu Gln His Ile Pro Val Phe Ala Val Ser Thr Arg  
210 215 220

Leu Phe Asn Ser Ile Thr Ser Val Phe Tyr Gly Arg Val Leu Lys Phe  
225 230 235 240

Asp Glu Val Ala Arg Phe Leu Val Ser Tyr Gln His Trp Thr Tyr Tyr  
245 250 255

Pro Val Met Ile Phe Gly Arg Val Asn Leu Phe Ile Gln Thr Phe Leu  
260 265 270

Leu Leu Leu Thr Arg Arg Asp Val Pro Asp Arg Ala Leu Asn Leu Met  
275 280 285

Gly Ile Ala Val Phe Trp Thr Trp Phe Pro Leu Phe Val Ser Cys Leu  
290 295 300

Pro Asn Trp Pro Glu Arg Phe Gly Phe Val Leu Ile Ser Phe Ala Val  
305 310 315 320

Thr Ala Ile Gln His Val Gln Phe Thr Leu Asn His Phe Ser Gly Asp  
325 330 335

Thr Tyr Val Gly Pro Pro Lys Gly Asp Asn Trp Phe Glu Lys Gln Thr  
340 345 350

Lys Gly Thr Ile Asp Ile Thr Cys Pro Pro Trp Met Asp Trp Phe Phe  
355 360 365

Gly Gly Leu Gln Phe Gln Leu Glu His His Leu Phe Pro Arg Leu Pro  
370 375 380

Arg Gly Gln Leu Arg Lys Ile Ala Pro Leu Ala Arg Asp Leu Cys Lys  
385 390 395 400

Lys His Gly Met Pro Tyr Arg Ser Phe Gly Phe Trp Asp Asp Ala Asn  
405 410 415

Val Arg Thr Ile Arg Thr Leu Arg Asp Ala Ala Val Gln Ala Arg Asp  
420 425 430

Leu Asn Ser Ala Pro Cys Pro Lys Lys Leu Gly Tyr Gly Glu Ala Tyr  
435 440 445

Asn Thr His Gly \*

450